



National Database for Autism Research

Strategies for Sharing Heterogeneous Data

February 6, 2012



Greg Farber, Director
Office of Technology Development and Coordination
National Institute of Mental Health

Take Home Messages

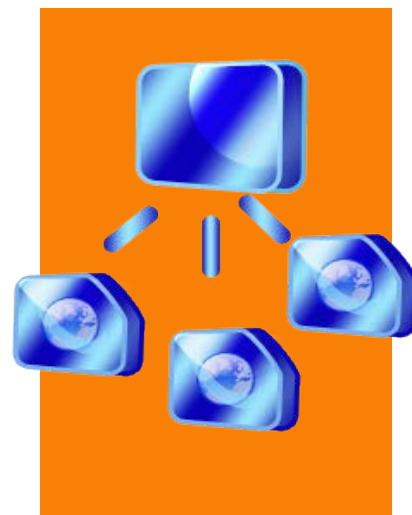
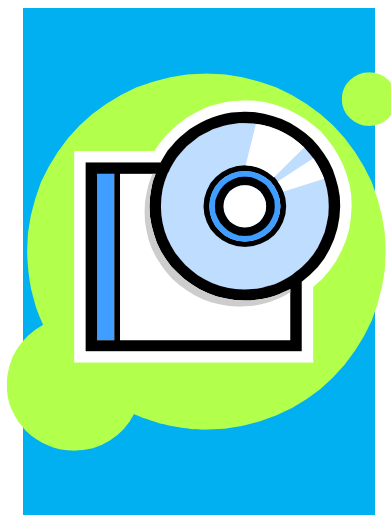
- The maturity of a field will determine the data sharing mechanisms that have a chance to work.
- Young fields often need a “nudge” from either the government, the journals, or both
- If there is any possibility of a subject participating in experiments in different labs, a GUID should be used.
- If the data sharing infrastructure is not easy for the investigators to use, they won't.

National Database for Autism Research (NDAR)

- Joint initiative supported by NIMH, NICHD, NINDS, and NIEHS and the NIH Center for Information Technology
 - Federally controlled system
 - Organizationally NDAR resides in the NIMH Office of Technology Development and Coordination
 - Data access committee is made up of NIH Program Officers/Human Subjects Experts
 - Trans-NIH program officer advisory committee
 - Relevant ICDs act as the Board of Governors for the project
- Begun in late 2006
- First data received in 2008 (NIH ACEs)

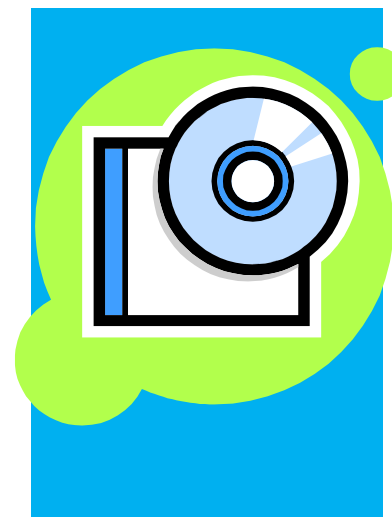
NDAR Implementation:

- **Global Unique Identifier (GUID)** – universal subject ID
- **Data Dictionary** – flexible framework for harmonizing data across studies and institutions
- **Data Sharing Regimen** – community and NIH input
- **Federation** – deep linkage to other relevant repositories



NDAR Data Dictionary

- Flexible and extensible framework for data definition
- 200+ instruments, image definitions, genomics defined
 - 29,000+ unique data elements and growing
 - A research community platform for defining the complex language characterizing autism research
 - Clinical
 - Genomics/Proteomics
 - Imaging Modalities
 - Eye Tracking
- Accommodates any data type and data structure
- Extended and enhanced by the ASD research community
- Curated by NDAR



Funded by the National Institutes of Health, the National Database for Autism Research (NDAR) is a secure research data repository promoting scientific data sharing and collaboration among autism spectrum disorder (ASD) investigators. NDAR's goal is to help accelerate scientific discovery in ASD through data sharing, data harmonization, and the reporting of research results. Data from over 25,000 research participants are available to qualified investigators through the NDAR portal. Summary information on the research projects sharing data, the type of data shared and data definitions are provided below.

[Request Access](#)

[Contribute](#)

[Preview Data](#)

[Research Spotlight](#)

[NDAR Data Dictionary](#)

[Autism Publications](#)

Available Records: 106,210

Select Data Source(s):

[Summary Data Approach - pdf](#)

☒ NDAR

☒ Pediatric MRI Data Repository

Select Gender

Select Phenotype

Select Sub Phenotype

Select Phenotype
Autism-like Developmental Disorders
Fragile X
Non-ASD Control
Not Enough Data Provided

Select Sub Phenotype
Select Sub Phenotype
Mildly Affected
Affected
Severely Affected

Age From (in months)

Select Non-Verbal IQ

Available Data Type(s)

Clinical Assessments

Aberrant Behavior Checklist - Community	544
Autism Diagnostic Interview - Revised (ADI-R) - (2003)	1578
Autism Diagnostic Interview-Rev (ADI-R) Toddler 2004	47
Autism Diagnostic Interview-Rev (ADI-R) Toddler 2006	47
Autism Diagnostic Observation Schedule (ADOS) Toddler	704
Autism Diagnostic Observation Schedule (ADOS)-Module 1	811
Autism Diagnostic Observation Schedule (ADOS)-Module 2	403
Autism Diagnostic Observation Schedule (ADOS)-Module 3	780
Autism Diagnostic Observation Schedule (ADOS)-Module 4	264
Autism Quotient Adult	12
CELF-4	45
CHARGE Family Characteristics Questionnaire	627

Imaging

MR Spectroscopy	1582
MRI DTI	124
MRI Structural	1601
fMRI	29

Genomics

Drug metabolism analysis / microarray	
Gene expression / PCR array	
Gene expression / microarray	3
Gene expression / sequencing	
Gene regulation / PCR array	
Gene regulation / mass spectrometry	
Gene regulation / microarray	820
Gene regulation / sequencing	
Metabolomic analysis / chromatography	
Metabolomic analysis / mass spectrometry	
Protein interaction / biosensor systems	
Protein interaction / immunoassay	



Preview Data

Research Spotlight

NDAR Data Dictionary

Autism Publications

[Link to Data](#)

Collection Title: Biological and Information Processing Mechanisms Underlying Autism NDAR

Investigators: Nancy Minshew, M.D. Mark Strauss, Ph.D. Kevin Pelphrey, Ph.D. Marcel Just, Ph.D. Thomas Mitchell, Ph.D. Diane Williams, Ph.D.
(Owner: Minshew, Nancy)

Collection Description: This center focuses on elucidating fundamental information processing and neurobiological mechanisms causing autism with studies of infant siblings, first-diagnosed toddlers, and groups of children, adolescents, and adults with and without

[More: \(Shared Summary Data, 14 Publications\)](#)

Grant Information:

Project Number	Project Title	Project Start Date	Project End Date	Organization Name
P50HD55748	Biological and Information Processing Mechanisms Underlying Autism	08/06/2007	07/31/2012	UNIVERSITY OF PITTSBURGH AT PITTSBURGH

[Show](#)

Casanova, Manuel F; El-Baz, Ayman; Elnakib, Ahmed; Switala, Andrew E; Williams, Emily L; Williams, Diane L; Minshew, Nancy J; Conturo, Thomas E, "Autism : the international journal of research and practice," *Quantitative analysis of the shape of the corpus callosum in patients with autism and comparison individuals*.

Conturo, Thomas E; Williams, Diane L; Smith, Charles D; Gultepe, Eren; Akbudak, Erbil; Minshew, Nancy J, "Journal of the International Neuropsychological Society : JINS," *Neuronal fiber pathway abnormalities in autism: an initial MRI diffusion tensor tracking study of hippocampo-fusiform and amygdalo-fusiform pathways*.

Damarla, Saudamini Roy; Keller, Timothy A; Kana, Rajesh K; Cherkassky, Vladimir L; Williams, Diane L; Minshew, Nancy J; Just, Marcel Adam, "Autism research : official journal of the International Society for Autism Research," *Cortical underconnectivity coupled with preserved visuospatial cognition in autism: Evidence from an fMRI study of an embedded figures task*.

{14} publications found, displaying 3 record(s)

[Publications](#)

Data Structures:

Title	Type	Records	Unique Subjects
Autism Diagnostic Interview - R (ADI-R) - (2003) v2	Clinical Assessments	144	144
Autism Diagnostic Observation Schedule - Module 1 v1	Clinical Assessments	58	58
Autism Diagnostic Observation Schedule - Module 2 v1	Clinical Assessments	24	24
Autism Diagnostic Observation Schedule - Module 3 v1	Clinical Assessments	76	76
Autism Diagnostic Observation Schedule - Module 4 v1	Clinical Assessments	138	138
CHARGE Family Characteristics Questionnaire v1	Clinical Assessments	286	286
CHARGE Medical History (May 2007) v1	Clinical Assessments	229	229
CHARGE Physical Exam (rev July 2007) v1	Clinical Assessments	223	223
M-CHAT v1	Clinical Assessments	44	44
Modified CHARGE Family Medical History (rev July 2007) v1	Clinical Assessments	286	286
Mullen - Scales of Early Learning v1	Clinical Assessments	86	86
Social Communication Questionnaire (SCQ) - Lifetime v1	Clinical Assessments	177	177

[Summary Data](#)

NDAR Data Dictionary

[Preview Data](#)
[Research Spotlight](#)
[NDAR Data Dictionary](#)
[Autism Publications](#)

[Clinical Assessments](#) | [CPEA STAA](#)

Clinical Assessments

- [Aberrant Behavior Checklist - Community.v01](#)
- [AGRE ADOS Module 1 2001.v01](#)
- [AGRE ADOS Module 2 2001.v01](#)
- [AGRE ADOS Module 3 2001.v01](#)
- [AGRE ADOS Module 4 2001.v01](#)
- [AGRE ADOS-G Module 1 \(2000 or earlier\).v01](#)
- [AGRE ADOS-G Module 2 \(2000 or earlier\).v01](#)
- [AGRE ADOS-G Module 3 \(2000 or earlier\).v01](#)
- [AGRE ADOS-G Module 4 \(2000 or earlier\).v01](#)
- [AGRE Autism Diagnostic Interview - R \(ADI-R\)](#)
- [AGRE Autism Diagnostic Interview - R \(ADI-R\)](#)
- [ATP Brain Donor Genetic Tests.v01](#)
- [ATP Brain Donor Genetic Tests.v01](#)

[Home](#)
[About NDAR](#)
[Data Sharing](#)
[Standards](#)
[Tools](#)
[Policy & Procedures](#)
[Training](#)
[FAQs](#)

Back to: [Home](#) > [Data Structures](#)

AGRE ADOS Module 1 2001

[Download Definitions](#)
[Download CSV Template](#)

Detail Definition

ElementName	Data Type	Size	Required	ElementDescription	ValueRange
subjectkey	GUID	0	Required	The NDAR Global Unique Identifier (GUID) for subjects which identifies a subject in NDAR	NDAR*
src_subject_id	String	20	Required	The site's subject identification	
src_record_id	Integer	0	Optional	Research site ID for record	
interview_date	Date	0	Required	Date on which the interview/genetic test/sampling/imaging was completed	MM/yyyy; MM/dd/yyyy
interview_age	Integer	0	Required	Age in months at the time of the interview/test/sampling/imaging.	0 to 1200
gender	String	50	Optional	Gender	M; F
untestable	String	50	Optional	Was the patient untestable?	Yes; No; N/A
untestable_reason	String	255	Optional	what was the reason he/she was untestable?	
olang	Integer	0	Optional	Overall Level of Non-echoed Language	0; 1; 2; 3; 8
fvoc	Integer	0	Optional	Frequency of Vocalization Directed to Others	0; 1; 2; 3
inton	Integer	0	Optional	Intonation of Vocalizations or Verbalizations	0; 1; 2; 8



The **REDCap Consortium** is comprised of **307 active institutional partners** from CTSA, GCRC, RCMI and other institutions, and it supports a secure web application (REDCap) designed exclusively to support data capture for research studies. The REDCap application allows users to build and manage online surveys and databases quickly and securely, and is currently in production use or development build-status for more than **25,010 studies** with over **36,140 end-users** spanning numerous research focus areas across the consortium.

Preview Data

Research Spotlight

NDAR Data Dictionary

Autism Publications

Link to Publication

☒ Articles with associated collections ☐ All Articles



32 records found, displaying 32 records(s), from 1 to 32. Page 1 / 1.

Title

[Social effects of oxytocin in humans: context and person matter.](#)

Bartz JA, Zaki J, Bolger N, Ochsner KN

Mount Sinai School of Medicine, Seaver Autism Center, Department of Psychiatry, New York, New York, 10029, USA. jennifer.bartz@mssm.edu

(PMID:21696997)

[Diffusion based abnormality markers of pathology: toward learned diagnostic prediction of ASD.](#)

Ingalhalikar M, Parker D, Bloy L, Roberts TP, Verma R

Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, USA.

(PMID:21609768)

[Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations.](#)

O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S, Karakoc E, Mackenzie AP, Ng SB, Baker C, Rieder MJ, Nickerson DA, Bernier R, Fisher SE, Shendure J, Eichler EE

Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington, USA.

(PMID:21572417)

Journal

Trends Cogn Sci. 2011 Jul;15(7):301-9

NDAR

Neuroimage. 2011 Aug;57(3):918-27

NDAR

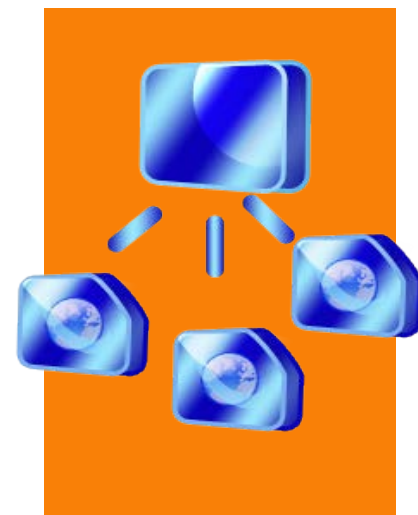
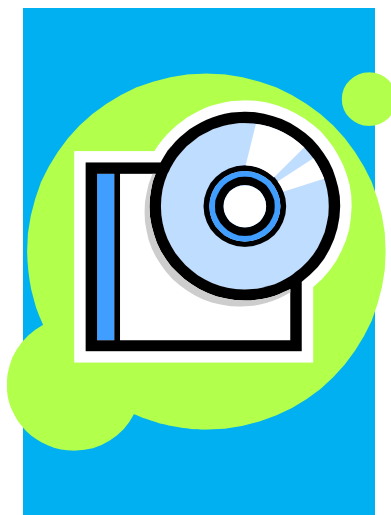
Nat Genet. 2011 Jun;43(6):585-9

NDAR

Link to Data

NDAR Implementation:

- **Global Unique Identifier (GUID)** – universal subject ID
- **Data Dictionary** – flexible framework for harmonizing data across studies and institutions
- **Data Sharing Regimen** – community and NIH input
- **Federation** – deep linkage to other relevant repositories



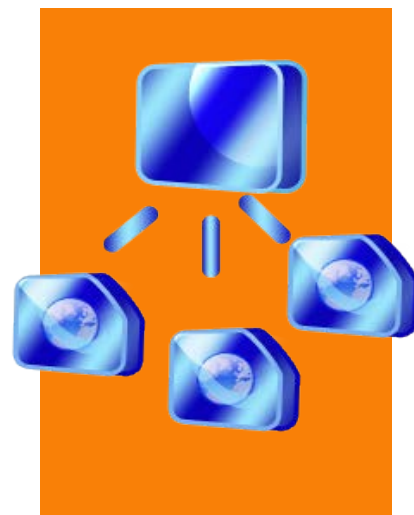
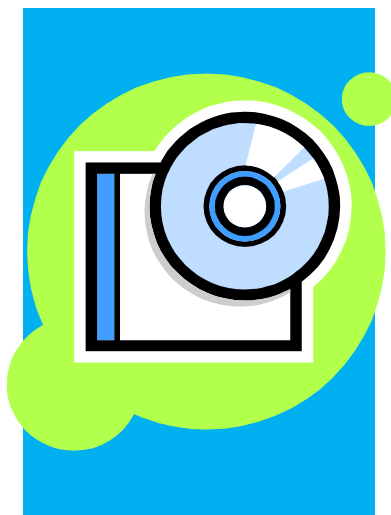
Data Sharing Regime

- Policies, procedures, & agreements
 - [NDAR Policy](#)
 - [Data Submission Agreement](#)
 - [Data Access Agreement](#)
 - [NDAR Standard Operating Procedures](#)
 - NDAR Data Federation MOU
- Data sharing expectations in FOAs
 - [PA-10-158](#) R01s for ASD Research
 - [PA-10-159](#) R03s for ASD Research
 - [PA-10-160](#) R21s for ASD Research
 - Other RFAs
- [Cost model](#) for data sharing expected with applications
- [Data sharing terms](#) included in the terms of award
 - Descriptive (clinical, baseline, structural MRI, raw genomics/proteomics)
 - Submitted every six months – shared 4 months later
 - Experimental (all other non-identifying data)
 - End of grant + 1 year or publication of primary objectives



NDAR Implementation:

- **Global Unique Identifier (GUID)** – universal subject ID
- **Data Dictionary** – flexible framework for harmonizing data across studies and institutions
- **Data Sharing Regimen** – community and NIH input
- **Federation** – deep linkage to other relevant repositories



Koser, Brian

NDAR Message Center

No New Messages

Submit Data

Query Data

User Dashboard | [Admin Dashboard](#)

My Access	My Collections	My Studies	My Data Structures	My Queries	Submissions
-----------	----------------	------------	--------------------	------------	-------------

The National Database for Autism Research



Query Data

Peditric MRI Study of Normal Brain Development



Query Data

ATP Federated Clinical Assessments



Apply for Access

ATP Brain MRI data and Images



Apply for Access

IAN Interactive Autism Network



Query Data

AGRE Restricted Access Permissions Group



Apply for Access

AGRE General Access Permission Group



Apply for Access

Take Home Messages

- The maturity of a field will determine the data sharing mechanisms that have a chance to work.
- Young fields often need a “nudge” from either the government, the journals, or both
- If there is any possibility of a subject participating in experiments in different labs, a GUID should be used.
- If the data sharing infrastructure is not easy for the investigators to use, they won't.

Contact Information

- Web site: <http://ndar.nih.gov>
- Email: ndarhelp@mail.nih.gov or halldan@mail.nih.gov
- Watch Short Video Tutorials at:
<http://ndar.nih.gov/ndarpublicweb/training.go>